

(43) Date of A Publication 06.10.1999

(21) Application No 9807264.8

(22) Date of Filing 03.04.1998

(71) Applicant(s)  
**3Com Technologies**  
(Incorporated in the Cayman Islands)  
PO Box 309, Upland House, Georgetown,  
Grand Cayman, Cayman Islands

(72) Inventor(s)  
**Paul James Moran**  
**David Law**  
**Peter Wilson**

(74) Agent and/or Address for Service  
**A A Thornton & Co**  
Northumberland House, 303-306 High Holborn,  
LONDON, WC1V 7LE, United Kingdom

(51) INT CL<sup>6</sup>  
**H04L 12/56 12/44**

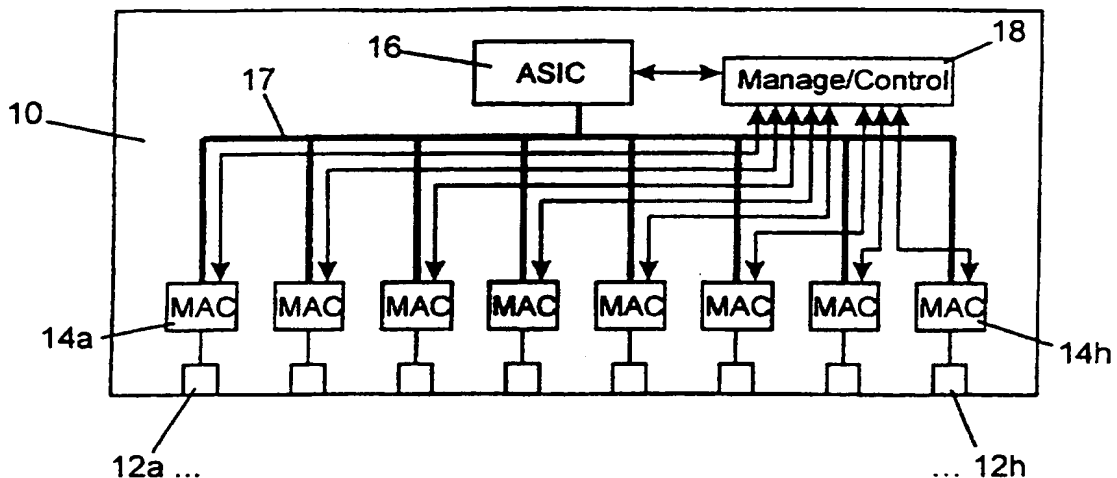
(52) UK CL (Edition Q )  
**H4P PPBC PPEC**

(56) Documents Cited  
**GB 2280574 A** **EP 0753950 A1** **WO 98/09408 A1**  
**US 5604867 A**

(58) Field of Search  
UK CL (Edition P ) **H4P PFD PPBC PPEC PPS**  
INT CL<sup>6</sup> **H04L 12/44 12/56**  
Online : **WPI**

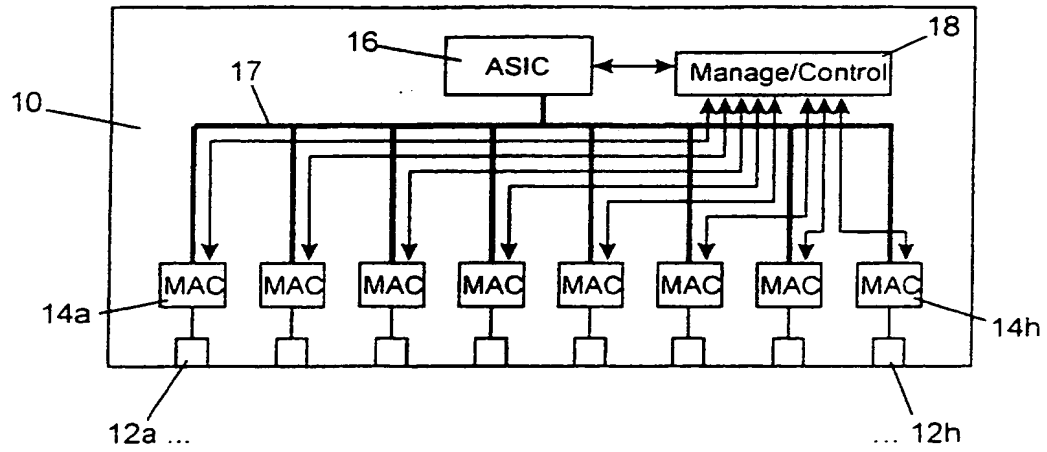
(54) Abstract Title  
**Controlling data flow through the ports of a communications device**

(57) A communication device (bridge 10) for use in a computer network is provided, the various ports 12e-12h of which may be configured to function at different maximum rates. This permits maximum flexibility in the configuration of the network and allows a network manager to take account of the different needs of various users of the network when the configuration takes place. A counter (20), Fig.2, (not shown), provides a "leaky bucket" implementation of the flow control for each port.

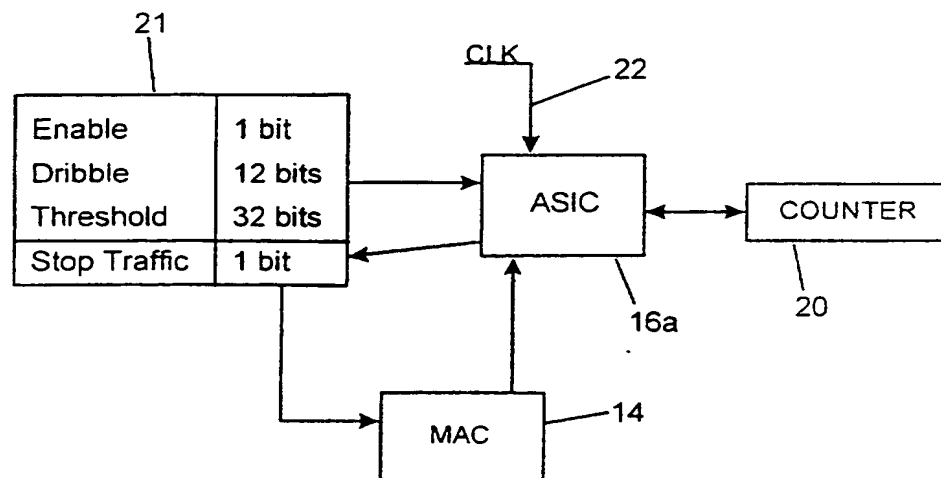


**Fig. 1**

1/2



**Fig. 1**



**Fig. 2**

**Fig. 3**

dribble clock /Hz					15258		
Possible Mbps	Req. Mbps	fill rate/sec	ideal dribble rate	subtract value	actual dribble rate	actual Mbps	
10	1	1250000	125000		9	137322	1.10
10	2	1250000	250000		17	259386	2.08
10	3	1250000	375000		25	381450	3.05
10	4	1250000	500000		33	503514	4.03
10	5	1250000	625000		41	625578	5.00
10	6	1250000	750000		50	762900	6.10
10	7	1250000	875000		58	884964	7.08
10	8	1250000	1000000		66	1007028	8.06
10	9	1250000	1125000		74	1129092	9.03
10	10	1250000	1250000		82	1251156	10.01
100	1	12500000	125000		9	137322	1.10
100	2	12500000	250000		17	259386	2.08
100	3	12500000	375000		25	381450	3.05
100	4	12500000	500000		33	503514	4.03
100	5	12500000	625000		41	625578	5.00
100	6	12500000	750000		50	762900	6.10
100	7	12500000	875000		58	884964	7.08
100	8	12500000	1000000		66	1007028	8.06
100	9	12500000	1125000		74	1129092	9.03
100	10	12500000	1250000		82	1251156	10.01
100	15	12500000	1875000		123	1876734	15.01
100	20	12500000	2500000		164	2502312	20.02
100	25	12500000	3125000		205	3127890	25.02
100	30	12500000	3750000		246	3753468	30.03
100	35	12500000	4375000		287	4379046	35.03
100	40	12500000	5000000		328	5004624	40.04
100	45	12500000	5625000		369	5630202	45.04
100	50	12500000	6250000		410	6255780	50.05
100	55	12500000	6875000		451	6881358	55.05
100	60	12500000	7500000		492	7506936	60.06
100	65	12500000	8125000		533	8132514	65.06
100	70	12500000	8750000		574	8758092	70.06
100	75	12500000	9375000		615	9383670	75.07
100	80	12500000	10000000		656	10009248	80.07
100	85	12500000	10625000		697	10634826	85.08
100	90	12500000	11250000		738	11260404	90.08
100	95	12500000	11875000		779	11885982	95.09
100	100	12500000	12500000		820	12511560	100.09

Network Device

5 The present invention relates to the control of communication traffic in a computer network and in particular relates to an arrangement in which the volume of traffic on a network can be controlled.

10 As is well known computer networks are formed of interconnected network devices such as file servers, printers, workstations and PCs. In a typical configuration, these network devices are interconnected in a star or tree formation by using communication hubs to which the various network devices are connected. Each communications hub has a plurality of ports to which either a further hub or a network device can be connected. Communication hubs retransmit communications received on any one of their ports to others of their ports to enable the communications sent by the network devices to reach their intended destinations. Typically, for instance in the Ethernet standard networks, communications occur by way of data packets which identify the source and destination addresses of the data packet in question as well as carrying the actual data to be communicated.

15 There are in existence a number of types of communication hubs, for instance repeaters, bridges and switches. These various types of devices handle the communications they receive in different manners, for instance a repeater simply retransmits all communications received on any port to all others of its ports, while a switch or bridge determines the intended destination of a received communication packet and, if it knows the intended destination, retransmits that packet only on the port or ports necessary for it to reach its intended destination.

20 Repeaters, as mentioned above, have an essentially passive functionality, while bridges and switches are more complex devices in which some control over the data passing through the device is exercised. The present invention is applicable in devices, such as bridges or switches, which exercise such control.

Over the past several years, developments have occurred in the implementation of computer networks and in particular in the Ethernet standard communications in networks. In particular, the permitted maximum speeds of communication which can be achieved have increased and, while a few years ago a  
5 typical network could carry communications at a rate of 10 megabits per second (Mbps) there are now in existence standards which permit communications to occur at speeds of up to 100 Mbps or 1000 Mbps.

To allow inter-operability between network devices designed to operate at different speeds, there are proposed communication hubs which can handle data at  
10 two or three different speeds, for instance both 10 and 100 Mbps, or 10, 100 and 100 Mbps.

It has been found however that there is resistance, for instance from network administrators, to installing highspeed networks in which all users of the network have access at the maximum possible rate. Reasons for this include the fact  
15 that if all users on the network were able to access the network at the higher rates, this would dramatically increase the load on the existing file servers and other pieces of network equipment, for instance those designed to interface with wide area networks. Such an increase in load may in fact reduce the overall capacity of the network, contrary to the intention of installing a fast network.

In this context therefore the present invention provides a  
20 communications hub for a computer network comprising a plurality of ports to which network devices may be attached and via which communications passing to and from the respective network devices can pass, comprising means associated with at least one of said ports arranged to set a limit for the rate at which communications can  
25 occur through the port said limit being less than the theoretical maximum communications rate via that port.

Such a communication device enables a network manager to

individually set a limit on the communication rate at which each user of the network can operate thereby preventing or reducing the risk of the network becoming overloaded as outlined above. Different users may be permitted to access the network at different communication speeds according to the requirements of the different users and their relative priorities for access to the network.

One implementation of a communication device as outlined above would be for the port devices to allow communications to occur via the port only for a proportion of the time, equal to the proportion which the limit is of the maximum bandwidth which is allowed via the port. For instance, if in a 100 Mbps capacity network a particular user is to be allowed to communicate only at 20 Mbps, the port would be enabled for 20% of the time.

This approach does result in an accurate limitation in the bandwidth to the specified allowed level and for requests and responses which are completed within the permitted timeslot the response time is very good. However, this implementation results in a very bursty pattern of traffic on the network and also no distinction is made between light users and heavy users of the network. In particular, a light user of the network who never wishes to exceed its limit still has periods of time when it cannot access the network.

With these difficulties in mind, in a particularly preferred implementation of the present invention the control means associated with each port of the communication device has a counter. This counter is incremented in one direction according to the amount of data passing through the port, and is incremented in the other direction periodically by an amount representing the permitted maximum data rate allowed through the port.

So long as a user of the port does not exceed the maximum allowed bandwidth the counter remains at or just above zero and the user has continuous access to the network at the permitted rate. In addition, a threshold is set representing

a significant deviation from the maximum allowed communication rate, and if the counter reaches this threshold then further communication via the port is more severely limited until such time has elapsed at the time average of the communications via the port falls to or below the maximum permitted communication rate for the port.

It is necessary to allow some temporary deviation from the maximum permitted rate in normal operation because the network components are capable of handling transmissions up to the maximum network rate and thus the rated value of the network will still dictate the instantaneous speed at which data can arrive.

This implementation provides advantages in that a light user of the network has effectively permanent access to the network at the allowed communications speed and also the adjustment of the allowed data rates on different ports is easily effected simply by adjusting the amount which is periodically incremented in the counter.

The present invention will be better understood from the following description of a preferred embodiment given by way of example and for illustration purposes and with reference to the accompanying drawings, in which:

Figure 1 is a schematic illustration of a repeater device in which the invention is implemented;

Figure 2 is a schematic illustration of parts of the communicating device implementing the present invention; and

Figure 3 is a table of values used in particular implementations of the invention.

In general terms what is provided by this invention and illustrated in more detail in the drawings discussed below is a communication device for use in a computer network the various ports of which may be configured to function at different maximum rates. This permits maximum flexibility in the configuration of

the network and allows a network manager to take account of the different needs of various users of the network when the configuration takes place.

Figure 1 illustrates, in schematic form, some basic features of a repeater device according to a preferred embodiment of this invention.

5 In Figure 1 bridge device 10 comprises a plurality of ports 12a to 12h to which network devices may be connected in the normal fashion. Eight ports are shown in this embodiment by way of example, it being understood that it is quite common for a considerably large number of ports to be provided on a network device and this could equally well be done within the context of the present invention. 10 Associated with each port 12 is a corresponding Media Access Control (MAC) device and MAC devices 14a to 14h are illustrated associated with ports 12a to 12h respectively. As is well known, MAC devices are provided in communication devices in computer networks to control the first interface between the network medium connected to the port and the internal components of the communication device. The 15 basic operations of MAC devices are well known and are not discussed here as, except to note that a MAC device is able to monitor transmission errors in communication packets and controls the flow of data through the port.

Within bridge device 10 is provided an ASIC device 16. ASIC device 16 provides the core functions of the bridge 10, that is it functions to receive 20 incoming data from all of the ports and to retransmit the received data via the ports according to known protocols and schemes of operation. ASIC 16 is therefore provided with connections to the MAC devices 14 and, in the embodiments illustrated in Figure 1, this is by way of a data bus 17 via which ASIC 16 can receive data from and transmit data to all of the MAC devices 14. In such an architecture within the 25 bridge device, there will also be present appropriate control signalling to enable and disable access to the bus by the various devices attached to it but the details of this are not shown in the present figures for clarity and because they are not necessary for

an understanding of the present invention. The bus architecture illustrated in Figure 1 is merely by way of example, and any other connection architecture to carry the data between ASIC 16 and the MAC devices 14 could equally well be implemented in the context of the present invention, for instances a standard star connection could be used in which each MAC device 14 is provided with an individual data link to ASIC 16.

Also it would be possible for the MAC devices to be integrated into the ASIC device such that a single chip provides the combined functions, although the parts are described separately herein for clarity.

The final major component of the bridge 10 illustrated in Figure 1 is the management/control means 18 which is provided to control the overall functioning of bridge device 10. While this is shown as a separate device from ASIC 16 it may in fact be implemented as an integral part of ASIC 16 or it may be implemented as illustrated as a separate device. In any event, management control means 18 is in communication with ASIC 16 to control and monitor its basic function in the normal way, and is also provided with means for communicating with each of MAC devices 14 to monitor and control their operation.

The present invention is implemented along the following lines within bridge device 10. For each port on which a maximum permitted data rate is to be set, the amount of data passing through the port is monitored and at each time a transmission or reception of data occurs, a number corresponding to the amount of data passing through the port is added to a counter. Additionally, as an ongoing process, another number is regularly subtracted from the counter, if the value in the counter is above zero.

The value which is periodically subtracted from the counter represents the maximum permitted data rate which is allowed through the relevant port and it will therefore be understood that, so long as the maximum permitted data rate is not

exceeded, the counter will not depart substantially from zero. If however, for a period of time, there is activity to the port considerably in excess of the maximum allowed bandwidth, the net value in the counter will continue to increase as an indication of this.

5                   Further in this implementation there is therefore set a threshold for the value in the counter and, if the counter value reaches the threshold value, then some flow control is implemented by the MAC on the port to restrict severely or prohibit further communication through the port until such time as the counter value has returned to a suitably low value.

10                  This flow control is according to known network protocols and, as is known, while it is possible immediately to disable further data transmission, it may be that further data will be received until the protocol for disabling the port from the network takes full effect. Such data is still received in the normal way, and is noted in the counter as normal.

15                  Such an implementation has been termed a leaky bucket implementation and it will be appreciated that the counter can be considered as a bucket for water having a hole in the bottom. Water can be poured into the top of the bucket and the size of the hole determines the rate at which water can dribble out. If a constant fill rate is less than or equal to the dribble rate then the bucket will always be empty,  
20                  while if a constant fill rate is greater than the dribble rate then the bucket will gradually fill up and eventually overflow.

                  If the fill rate is not constant but over time averages to be equal to or less than the dribble rate then, so long as any sudden influx is not greater than the capacity of the bucket, then the capacity of the bucket will allow all of the water to  
25                  dribble through the hole in the bucket without the bucket overflowing.

                  It will be appreciated that this is analogous to the data rate limitation implementation discussed above in which the leaky bucket is represented by the

counter, a non-zero value in the counter being equivalent to the presence of water in the bucket. Activity on the port is represented by adding water to the bucket and the periodic subtraction from the counter represents water dribbling out of the hole in the bucket. The bucket is set to sufficiently large that it can cope with the arrival of a reasonable number of maximum size communications packets effectively instantaneously, but the repeated arrival of such amounts of data may cause the bucket to overflow, depending on the dribble rate.

The threshold of the counter at which flow control for the port is implemented is analogous to the overflowing of the bucket which must be prevented if no water is to be lost.

Figure 2 is a schematic illustration of the various means implemented within bridge 10 concerned with the implementation of the present invention. In Figure 2 there are shown only the parts associated with a single port, it being understood that corresponding means are provided for each of the ports on which the flow control according to the present invention is implemented.

In Figure 2 there is provided a counter 20 which, in this implementation, is a 32 bit counter and represents the leaky bucket discussed above. Also shown in Figure 2 is a control means, in the form of a part of the ASIC designated 16a in Figure 2 which is responsible for the implementation of the present embodiment. Associated with the ASIC 16a is a set of registers 21 which hold the control values for the process according to the present embodiment.

In particular it will be seen that the register 21 include an enable value, a dribble value, a threshold value and a stop traffic value. The enable value simply indicates whether or not the flow control according to the this invention is to be implemented on the port. In this implementation this is a single bit and for the remainder of the discussion it will be assumed that this bit is set to indicate that the flow control should be implemented. The threshold value held in register 21 is a 32

bit value, corresponding to the fact that the counter 20 is a 32 bit counter, and sets the maximum value which will be allowed to accumulate in counter 20 before action is taken to reduce the flow. The dribble value held in register 21 is, in this implementation, a 12 bit value and this is the value which is periodically subtracted from the counter to represent the leak in the bucket. These values are set by the network control in order to set up the flow control for a particular port and these values are then input to the control implemented by the ASIC 16a.

The final value held in register 21 is the stop traffic value, and again this is a single bit value which is controlled by the ASIC and indicates to the MAC device 14 whether or not flow control currently requires implementation for the port in order to bring the traffic on that port within the maximum allowed level.

ASIC 16a also has input to it a clock signal on line 22 and this clock is termed the dribble clock and ASIC 16a acts once every clock cycle of the dribble clock to subtract the dribble value from register 21 from the counter 20. Finally, as illustrated in Figure 2, MAC 14 signals to ASIC 16a the quantity of data passing through the port in order for ASIC 16a to add the appropriate value to counter 20.

In the following there is described a particular numerical example of an implementation of this embodiment given by way of illustration. In this illustrative embodiment it will be the case that the numbers used to represent the quantity of data are actually the numbers of bytes (groups of 8 bits) passing through the port as this enables the use of smaller numbers. Figure 3 is a table setting out the various numbers which are used in this implementation of the embodiment.

In Figure 3 the first column, headed possible megabits per second, corresponds to the rated speed of the network. As shown, this is typically either 10 megabits per second or 100 megabits per second in accordance with known network implementations. We appreciated that, in the present invention, even when the control according to this invention is implemented to reduce the overall data speed

through a particular port, the network components are capable of handling transmissions up to the maximum network rate and therefore the rated value of the network speed will still dictate the instantaneous speed at which data can arrive. This therefore sets the fill rate in bytes per second which is shown in the third column of Figure 3 and it will be seen that the fill rate per second in bytes is simply  $1/8$  of the maximum value of megabits per second.

The second column of Figure 3 represents the various values which may be required to be set as the maximum allowed data rate through a particular port. Again, these are firstly given in the second column of Figure 3 in terms of megabits per second and these convert, again by factor of  $1/8$ , to the values given in the fourth column of Figure 3. This column is headed ideal dribble rate because, as mentioned above, the dribble rate, that is the rate at which numbers are subtracted from the counter, equates to the maximum allowed data rate through the port. The next column in Figure 3, headed subtract value, is dictated by the ideal dribble rate and the frequency of the dribble clock 22 in Figure 2. This is the value which will be subtracted for every clock cycle. In this embodiment, the frequency of the dribble clock is 15,258 Hz (derived from 125 MHz clock) and, broadly speaking, the subtract value in Figure 3 is the ideal dribble rate value divided by 15,258. The actual values are rounded to give whole values for storing in register 21 and for use in subtracting from the counter.

Finally, Figure 3 illustrates in its right most two columns the actual dribble rate and the actual maximum megabits per second which are allowed by the rounded subtract values, and it will be seen that these do not depart substantially from the ideal values given in the lefthand side of the table.

Taking, by way of example, the example from Figure 3 of a 100 Mbps network in which it is desired to restrict the data flow through a particular port to 25 Mbps. It will be understood then that the value of 205 is set in the dribble register is

register 21 in Figure 2. On each clock cycle of clock 22, operating at 15,258 Hz, this value is subtracted, by ASIC 16a, from the current value in counter 20 so long as the current value in counter 20 is not zero. Each time traffic is passed through the associated port, MAC device 14 communicates to ASIC 16a the number of bytes which have passed through the port and then this value is added to the value in counter 20. So long as the time averaged data flow through the port does not exceed 25 Mbps the value in counter 20 will not depart substantially from zero and data will continue to pass through the port.

Alternatively it may be that the value stored in the dribble register is in fact the 2's complement of the actual dribble value, such that the stored value can be added to the counter with an effect equivalent to subtracting the required dribble value.

As discussed above, there is also set in register 21 a threshold value and, after each addition to counter 20, ASIC 16a compares the new value in counter 20 with the threshold value from register 21. If the value in counter 20 is greater than the threshold value from register 21, ASIC 16a sets the stop traffic value to one. MAC device 14 is responsive to the value of the stop traffic bit such that when it is set to one MAC device 14 implements flow control through the port such that it immediately prevents any further transmission of data and, in accordance with known protocols, acts to prevent further reception of data.

ASIC 16a continues to subtract the dribble value from counter 20 on each clock cycle of clock 22 until such time as the value in counter 20 reaches zero. At this time the stop traffic bit is reset to zero, responsive to which the MAC device 14 again permits communication through the port.

In this implementation, if the current value in counter 20 is lower than the threshold value the value of the stop traffic bit is not changed. This means that the MAC device must wait until the counter has completely cleared until it is again

permitted to transmit traffic through the port.

In an alternative implementation it would be possible to provide a further intermediate threshold value at which the transmission of data could be recommenced. The setting of the threshold value or values defines how much  
5 departure from the ideal data rate will be permitted and is a matter for consideration by the network controller.

As mentioned above the threshold value may be set to permit the arrival of a reasonable number, perhaps one or two, of maximum size packets without triggering the flow control. If the threshold is set too low it may interfere with the  
10 reception and forwarding of packets to the detriment of the overall network. However, it is also important not to set the threshold too high because the higher the threshold, the longer the times will be during which communications will be prohibited or severely limited, while the counter increments to zero by the dribble clock.

15 It will therefore be seen that the present invention provides a simple and effective way in which quantity of data passing through ports in a communication device can be controlled, while also allowing for simple alteration of the control values simply by resetting the values stored in register 21.

CLAIMS:

1. A communications device for a computer network comprising a plurality of ports to which network devices may be attached and via which communications passing to and from the respective network devices can pass comprising control means associated with at least one of said ports arranged to set a limit for the rate at which communications can occur through the port said limit being less than the theoretical maximum communications rate via that port.
2. A communications device according to claim 1 comprising said control means associated with a plurality of ports in which said control means sets said limit for each of said plurality of ports independently of said limit set for the remainder of said plurality of ports.
3. A communications device according to claim 1 in which said control means allows each of said at least one port to carry communications only for a proportion of the time, said proportion corresponding to the ratio between said limit and said theoretical maximum rate.
4. A communications device for a computer network comprising a plurality of ports to which network devices may be attached and via which communications passing to and from the respective network devices can pass comprising, associated with at least one of said ports;
  - monitoring means for monitoring the amount of data passing through said port;
  - a counter means having a value therein;
  - means responsive to said monitoring means for adding to said counter

means values corresponding to said amount of data;

storage means for storing first and second predetermined values;

clock means providing a periodic clock signal having a predetermined clock interval;

5 means responsive to said clock signal for subtracting from said counter, at each said clock interval, the lower of said first predetermined value and the value in said counter;

means for determining if the value in said counter means exceeds said second predetermined value; and

10 means responsive to said means for determining for restricting the passage of communications through said port when said value in said counter means exceeds said second predetermined value.

5. A communications device according to claim 4 in which said means for restricting the passage of communications disables the transmission of data from the port until the value in said counter means falls to a third predetermined value.

6. A communications device according to claim 4 in which said means for restricting the passage of communications initiates a disabling of the reception of data at the port using a network protocol when said value in said counter means exceeds said second predetermined value and until the value in said counter means falls to a third predetermined value.

7. A communications device according to claim 5 or 6 in which said third predetermined value is zero.

8. A communications device according to claim 5 or 6 in which said third

predetermined value is a value intermediate zero and said second predetermined value.

- 5 9. A method of operating a communications device for a computer network comprising a plurality of ports to which network devices may be attached and via which communications passing to and from the respective network devices can pass;
- 10 the method comprising, for at least one of said ports;  
monitoring the amount of data passing through said port;  
monitoring a count value;  
adding to said count value corresponding to said amount of data;  
storing first and second predetermined values;  
providing a periodic clock signal having a predetermined clock interval;  
subtracting from said count value, at each said clock interval, the lower  
15 of said first predetermined value and the count value;  
determining if the count value exceeds said second predetermined value;  
and  
restricting the passage of communications through said port when said  
count value exceeds said second predetermined value.
- 20 10. A method according to claim 9 in which the transmission of data from the port is disabled when said count value exceeds said second predetermined value until the value in said counter means falls to a third predetermined value.
- 25 11. A communications device according to claim 9 in which the reception of data at the port is disabled using a network protocol when said value in said counter means exceeds said second predetermined value and until the value in said

counter means falls to a third predetermined value.

12. A communications device according to claim 10 or 11 in which said third predetermined value is zero.

5

13. A communications device according to claim 10 or 11 in which said third predetermined value is a value intermediate zero and said second predetermined value.



Application No: GB 9807264.8  
Claims searched: 1-13

Examiner: Keith Williams  
Date of search: 15 October 1998

**Patents Act 1977**  
**Search Report under Section 17**

**Databases searched:**

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:  
UK CI (Ed.P): H4P (PFD, PPBC, PPEC, PPS)  
Int CI (Ed.6): H04L 12/44, 12/56  
Other: Online WPI

**Documents considered to be relevant:**

Category	Identity of document and relevant passage		Relevant to claims
X	GB 2280574 A	Northern Telecom - see abstract; Claim 1	1,4,9
X	EP 0753950 A1	Compaq Computer Corp. - see column 3, lines 9-26 (and US 5742602)	1
X	WO 98/09408 A1	Advanced Micro Devices, Inc. - see page 15, line 10 to page 16, line 12	1,4,9
X	US 5604867	Network Peripherals - see Claim 1; column 21, lines 58 onwards; column 22, lines 66 onwards	1,4,9

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.